On the improvement of coevolutionary optimizers by learning variable interdependencies

Karsten Weicker
Institute of Computer Science, Dept. Formal Concepts
University of Stuttgart, Breitwiesenstr. 20–22
70565 Stuttgart, Germany
Karsten.Weicker@informatik.uni-stuttgart.de

Nicole Weicker
Institute of Computer Science, Dept. Formal Concepts
University of Stuttgart, Breitwiesenstr. 20–22
70565 Stuttgart, Germany
Nicole.Weicker@informatik.uni-stuttgart.de

Abstract-

During the last years, cooperating coevolutionary algorithms could improve convergence of several optimization benchmarks significantly by placing each dimension of the search space in its own subpopulation. Though, their general applicability is restricted by problems with epistatic links between problem dimensions – a major obstacle in cooperating coevolutionary function optimization. This work presents first preliminary studies on a technique to recognize epistatic links in problems and self-adapt the algorithm in such a way that populations with interrelated dimensions are merged to a common population.

1 Introduction

There are various, different types of genetic algorithm (GA) difficulty in optimization problems which complicate the solution of those problems by canonical standard GAs. Examples for GA difficulty are deceptiveness (Das & Whitley, 1991), multimodality (Deb, Horn, & Goldberg, 1993), and noise (Kargupta & Goldberg, 1994). A different type of GA difficulty is epistasis where random and highly ordered epistatic interactions are distinguished. In the first kind of epistasis a linkage between two genes has a random influence on the fitness like it is defined in the NK-landscape problem by Kauffman (1993). This work focuses on the second type of epistasis where there are interdependencies between the dimensions of the search space on the phenotype level. In this connection, Salomon (1996) generated epistatic problems by rotating standard benchmark functions which are mostly decomposable without rotation. Those functions are of special interest since they have a high degree of interdependencies which makes them very hard for standard GAs.

There are already various approaches to deal with epistasis in function optimization. For example, one rule of thumb is to place epistatic dimensions close together in the genotype. In order to do this, a-priori knowledge about the kind of epistasis is necessary. One approach to avoid this was presented by Paredis (1996) where a coevolutionary algorithm was used to evolve the solution as well as the arrangement of the search space dimensions.

This work chooses a different approach which modifies the coevolutionary method of Potter and De Jong (1994). This algorithm assigns each search space dimension to a separate subpopulation which are evolved each by its own GA. Interestingly, this method is suited for the optimization of many standard benchmarks. But particularly in the case of functions with a high degree of epistasis (cf. Potter, 1997) their performance falls behind the performance of canonical GAs.

The overall goal of the work in this article is to develop a method which can optimize epistatic as well as non-epistatic functions equally well. Nevertheless, the presented results are just first, preliminary studies. Consequently, they can only be a first step towards this goal.

The remainder of the paper is organized as follows. Section 2 discuses related work. The fundamental idea of the presented method is described in Section 3, and Section 4 presents the mechanism for the recognition of epistatic links. In Section 5 first results are discussed before Section 7 concludes and outlines future work.

2 Related work

Potter and De Jong (1994) proposed a cooperative coevolutionary algorithm which improves the performance of GAs on many benchmark functions significantly. This approach is discussed in more detail in (Potter, 1997) and has also been applied successfully to neural network design (Potter & De Jong, 1995), sequential decision rules (Potter, De Jong, & Grefenstette, 1995), and concept learning (Potter & De Jong, 1998). In the case of function optimization, each dimension of the search space is assigned to a population which is optimized by its own evolutionary algorithm. Those population perform each one generation one after another. For the fitness evaluation of an individual of one population partial solutions from all other populations are used and collaborate together for one solution of the complete problem. This species interaction is shown exemplary in Figure 1 for the evaluation of an individual of the first population. In the intuitive collaboration strategy the best individuals from the populations are used as representatives for the fitness evaluation.

This method could solve many benchmark problems with no or low degree of epistasis more efficiently, i.e. with less fitness evaluations, than canonical GAs (cf. Potter & De Jong, 1994). However, sometimes, the coevolutionary algorithm leads to worse results for functions with a high degree of epistasis, like those introduced by Salomon (1996). The

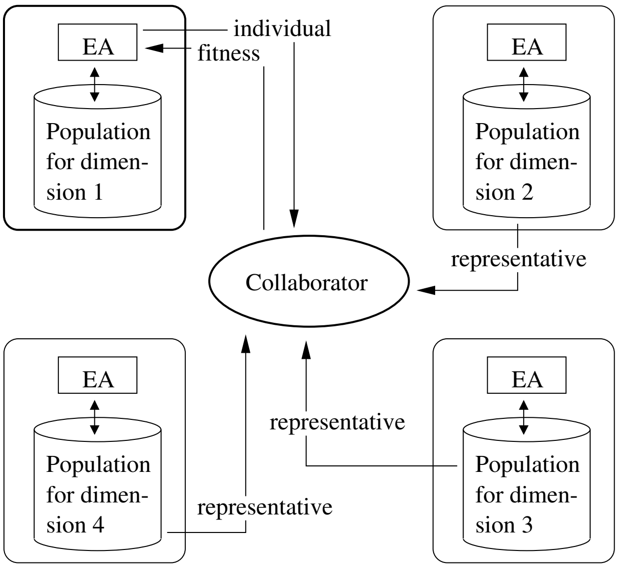

Figure 1: Species interaction in Potter approach

reason is that the described intuitive collaboration strategy is too greedy. Potter (1997) could reach a slight improvement with an alternative collaboration strategy where random individuals are chosen from all other populations. An individual collaborates with those randomly chosen individuals as well as the best individuals and gets the best fitness of both evaluations assigned.

On the other side there are first approaches to deal with epistasis. Thus Paredis (1996) has introduced a method which co-evolves the solution to the problem as well as the arrangement of the dimensions in the genotype. The benefit of this approach is obvious for problems with only little epistasis – the cited work regards problems where each dimension is linked with two other dimensions. Those problems enable a certain improvement by the method. However, for problems with full epistasis an improvement is scarcely possible since no arrangement has advantages over the other arrangements.

3 Adaptive coevolutionary optimization

This Section motivates the approach of this work and discusses the developed mechanism for cooperating coevolutionary algorithms.

Figure 2 shows two cases of the behavior of the canonical and the coevolutionary GA on a non-epistatic and an epistatic fitness function which may be found similar in (Potter, 1997). The left part shows the convergence on a benchmark function without epistasis and the right part on a rotated function with high epistasis. This is an extreme example where both the canonical GA and the coevolutionary GA are superior on different degrees of epistasis. In order to choose the best algorithm for a problem, again a priori knowledge on the epistasis is necessary.

Therefore, this example motivates the development of an optimization technique which adapts to the inherent epistasis of the problem and behaves roughly like the respective better algorithm. Note that the clearness of the characteristic in Figure 2 depends not only on the regarded fitness function but also highly on selection pressure and the type of selection.

In order to reach this goal the following method is proposed. First of all, it is assumed that there is no knowledge about the problem's epistasis and all dimensions are coevolved in seperate populations as proposed by Potter and De Jong (1994). If depending on certain criteria which are diskussed in Section 4 there is enough evidence that two search space dimensions j and k are interrelated, the populations

$$\begin{aligned} Pop_j &= \left\langle \, Ind_1^{(j)}, \ldots, Ind_n^{(j)} \, \right\rangle \\ Pop_k &= \left\langle \, Ind_1^{(k)}, \ldots, Ind_n^{(k)} \, \right\rangle \end{aligned}$$

are replaced by the merged population

$$Pop_{j,k} = \left\langle \, Ind_1^{(j,k)}, \ldots, Ind_n^{(j,k)} \, \right\rangle$$

with the individuals defined as

$$Ind_i^{(j,k)} = \begin{cases} Ind_{best_j}^{(j)} \circ Ind_{best_k}^{(k)}, & \text{for } i = 1 \\ Ind_{best_j}^{(j)} \circ Ind_{Uniform[1,n]}^{(k)}, & \text{for } 2 \leq i \leq \lceil \frac{n}{2} \rceil \\ Ind_{Uniform[1,n]}^{(j)} \circ Ind_{best_k}^{(k)}, & \text{otherwise} \end{cases}$$

where $best_j$ and $best_k$ are the indices of the respective best individuals, $\circ$ denotes the the concatenation of the bit-strings, and $Uniform[1, n]$ denotes a random number between 1 and n.

Thus, both dimensions are paced together in the same genotype and are evolved together. This approach is comparable to the aim of Paredis (1996) to place epistatic linked dimensions close together. But the presented merging technique preserves the advantage of evolving presumably independent dimensions in different species. The approach is shown exemplary in Figure 3 where an epistatic link between search space dimensions 1 and 3 is discovered. A new merged population is created and both single populations do not participate in the evolution anymore and are deleted.

Iteratively, the technique can create even bigger species containing more than two dimensions. In the extreme, this may lead to only one species for fully epistatic functions, which is equivalent to a canonical GA. Nevertheless, this preliminary study only regards that populations may contain only a maximum of two dimensions.

4 Recognizing epistatic links

The mechanism presented in the previous section depends on the successful recognition of epistatic links. Within the

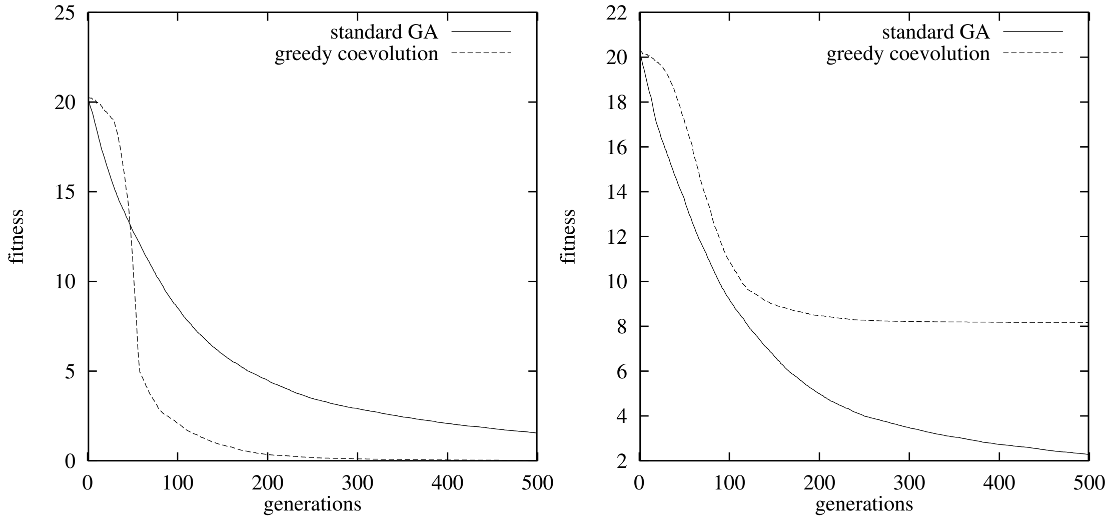

Figure 2: Performance of the canonical GA and the coevolutionary method with two different collaboration strategies on a non-epistatic (left) and an epistatic (right) fitness function (Ackley function and rotated Ackley function)

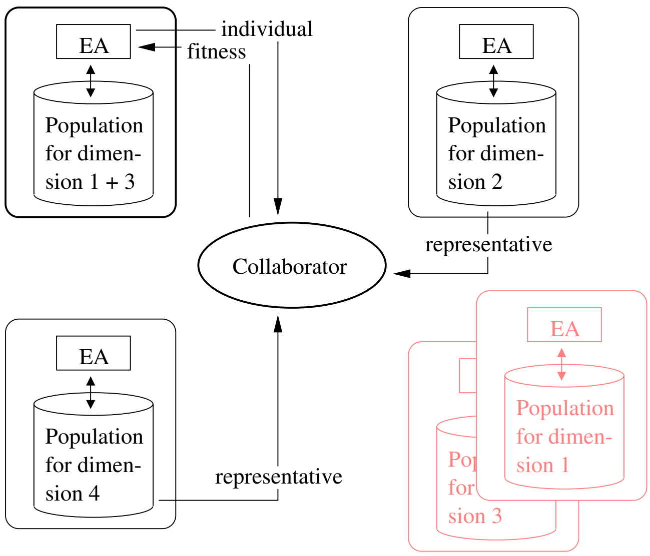

Figure 3: Species interaction where the populations for dimension 1 and 3 have been merged.

framework of this study various approaches have been tested and a promising technique is described in this section.

The technique relies on the observation that an indication for an epistatic link between two dimensions is given if a new candidate solution where both dimensions have been changed achieves a better fitness than a candidate solution where just one dimension was changed.

Therefore, similar to the non–greedy collaboration strategy of (Potter, 1997), a new individual for dimension i representing value $newval_i$ is evaluated by the collaborator in two different candidate solutions $cand_{best} = \langle d_1, \ldots, d_n \rangle$ and $cand_{random} = \langle d'_1, \ldots, d'_n \rangle$ which are defined by

$$d_j = \begin{cases} newval_i, & i = j \\ bestval_j, & \text{otherwise} \end{cases}$$

$$d'_j = \begin{cases} newval_i, & i = j \\ randval_k, & k = j \\ bestval_j, & \text{otherwise} \end{cases}$$

where $bestval_j$ is the value of the best individual in the population for dimension j, $randval_k$ is the value of a randomly chosen individual in the population for dimension k, and k is a randomly chosen dimension distinct from i.

However, the fitness values of the two candidate solutions are not used for the fitness evaluation of the individual of dimension j – like in the non-greedy collaboration strategy – but also for recognizing epistatic links. If $cand_{random}$ is better than $cand_{best}$ a counter for the link (j, k) is increased.

These counters are analyzed at the end of an evaluation round for each population. In an iterative process only those populations are merged where the counter for the link is the maximal occurring counter for both search dimensions under consideration of all not yet merged populations.

This technique has lead to the best results in our experiments. There have also been experiments with more sophisticated recognition mechanisms or complete evaluations rounds only for the learning of epistatic links. Nevertheless, the learning mechanism is very sensible with respect to two details.

First, the learning process should not start too early since the fitness level is too bad at the beginning and, therefore, many wrong counter increases causes merging of populations with non-epistatic dimensions.

Second, the calibration of the acceptance condition for epistatic links is essential for a successful learning process. So far no extensive systematical study was executed to de-

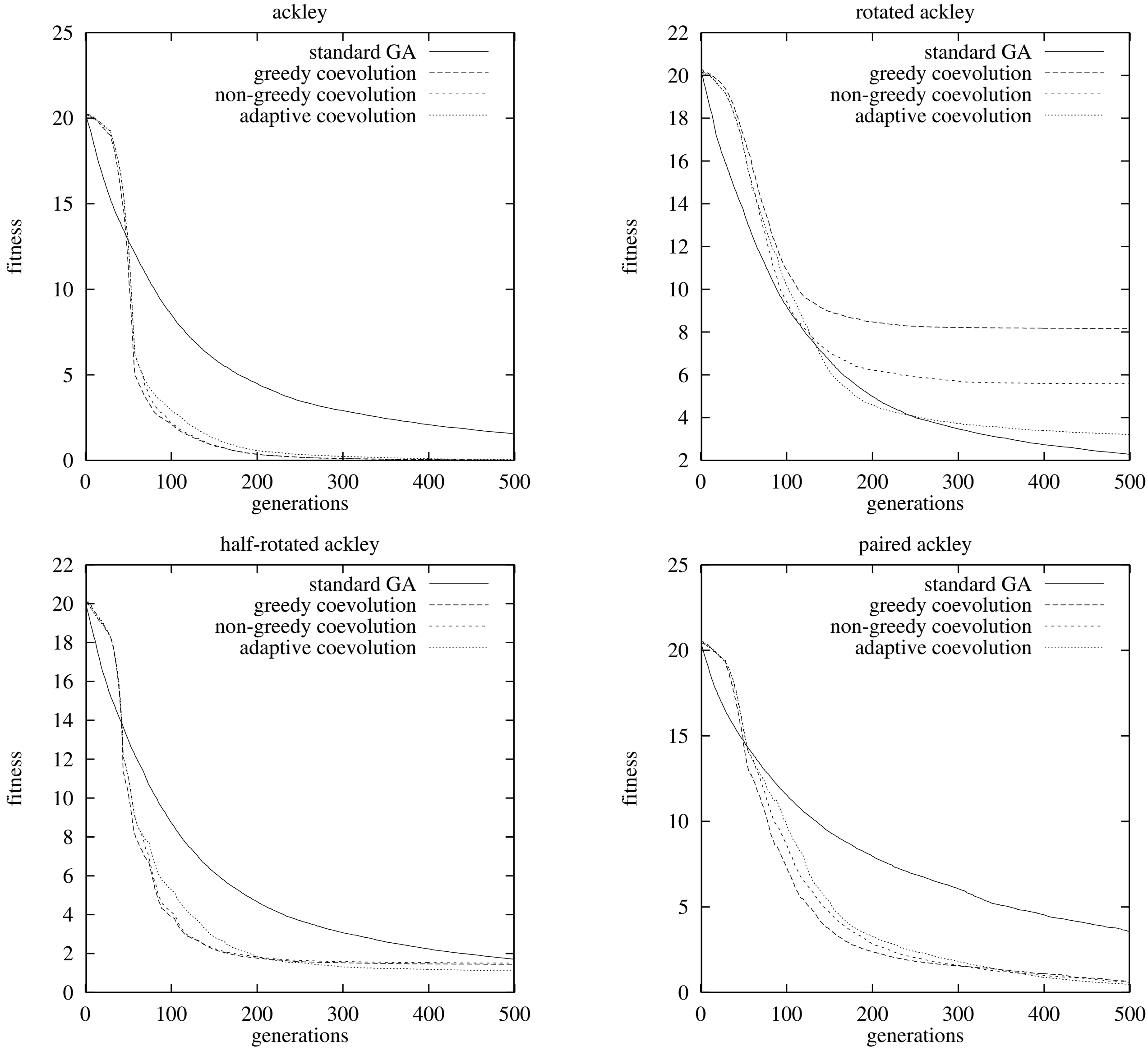

Figure 4: Convergence curves for various versions of the Ackley function.

termine an optimal acceptance strategy. For the existing experiments the above strategy was chosen which turned out to work to some extent. Future strategies should also consider the number of tests for the respective links and perform a more sophisticated analysis of the counter matrix.

The described strategy was combined with the non-greedy collaboration strategy, i.e. that for each evaluation three fitness values are computed where only two are used for the learning mechanism and all three for the individual's fitness.

5 Experiments

This Section presents first results with the presented adaptive coevolutionary algorithm. In order to compare the results of this technique with the results reported by Potter (1997), the sane parameter settings have been used: a population size of 100, a two-point crossover with rate 0.6, a bitflipping mutation with a rate reciprocal to the chromosome length, balanced linear scaling, and a selection where for each newly selected individual in an iterative process individuals are chosen uniformly until one is accepted with probability $fitness/(maximum - fitness)$. Note that this selection mechanism leads to different results than true fitness proportionate selection.

As test suite the generalized Ackley function (Bäck & Schwefel, 1993), the Rosenbrock function (Spedicato, 1975), and the Schwefel function (Schwefel, 1977) have been used with dimensionality 30. In addition random rotation technique by Salomon (1996) was applied to the test functions. On each test function a canonical GA, the coevolutionary al-

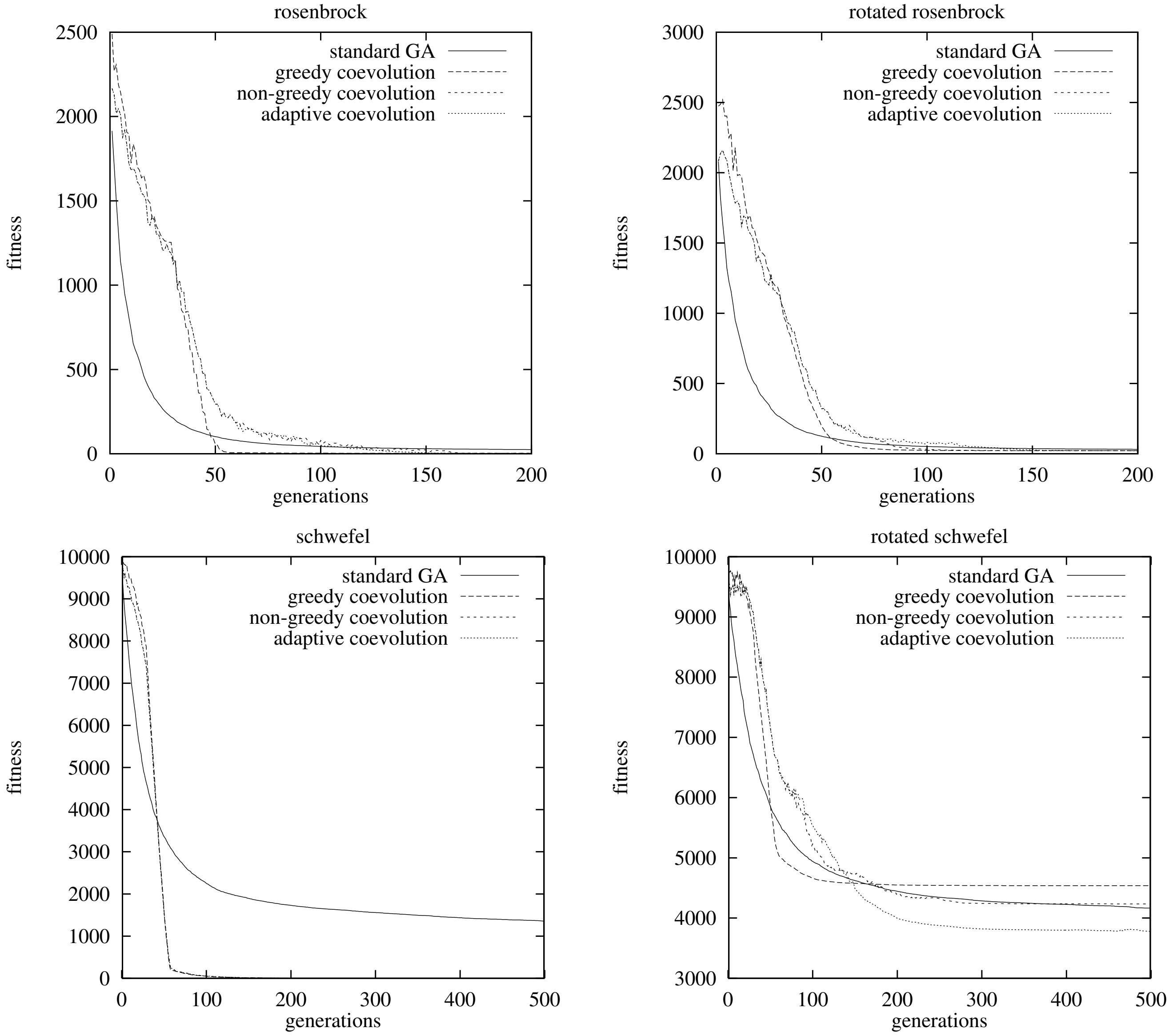

Figure 5: Convergence curves for Rosenbrock and Schwefel function.

gorithm of Potter and De Jong (1994) with the greedy as well as the non-greedy collaboration strategy, and the presented adaptive coevlutionary algorithm have been carried out for 50 experiments. The results shown in the diagrams are the best fitness values in the populations averaged over the different experiments.

Figure 4 shows the results for various versions of the Ackley function. Besides the standard and the rotated function, a half-rotated and a paired version is used. The half-rotated function consists of the sum of two 15-dimensional Ackley functions where only one function is rotated, i.e. that only 15 dimensions are fully epistatic linked. In the paired Ackley function the inputs for a non-rotated Ackley function are partitioned in pairs to which the transformation $(x, y) \mapsto (x, y - x)$ is applied. This establishes 15 epistatic links. The results show that in all cases the adaptive coevolutionary algorithm is able to approximate the respective best method. The trend in the case of the half-rotated Ackley function indicates that with further improvement of the merging technique in order to merge more than two populations the adaptive algorithm might outperform the other algorithms even under consideration of the additional fitness evaluations.

Figure 5 shows the results for the Rosenbrock and the Schwefel function. On both functions the algorithm shows a rather slow convergence at the beginning, but in the case of the latter function the algorithm is to outperform the other methods clearly.

6 Discussion and future work

The focus of the presented work is on the exploration of the basic potential of such an approach. There was neither a fine-tuning of the method nor an effort to minimize the additional fitness evaluations. Nevertheless, the preliminary studies show that it is worth to invest more work in the elaboration of the technique.

First, the recognition mechanism has to be improved since the present version inclines to merging populations which have been processed earlier. Moreover, more sophisticated acceptance conditions are necessary to create populations for more than two dimensions.

Second, experiments show that different kinds of selection or selection pressure in the same algorithm leads to completely different results. Actually, this underlines the necessity of such a self-adaptive algorithm – nevertheless it also indicates that the selection pressure could be adapted too in the algorithm.

7 Conclusion

The presented adaptive coevolutionary algorithm and its first results seem to be promising – especially under consideration of the provisional character of current algorithm. It is a first step towards the overcoming of the main obstacle in cooperating function optimization, namely epistatic links in optimization problems.

References

Bäck, T., & Schwefel, H.-P. (1993). An overview of evolutionary algorithms for parameter optimization. *Evolutionary Computation*, *1*(1), 1–23.

Das, R., & Whitley, D. L. (1991). The only challenging problems are deceptive: Global search by solving order-1 hyperplanes. In R. K. Belew & L. B. Booker (Eds.), *Proc. of the Fourth Int. Conf. on Genetic Algorithms* (pp. 166–173). San Mateo, CA: Morgan Kaufmann.

Deb, K., Horn, J., & Goldberg, D. E. (1993). Multimodal deceptive functions. *Complex Systems*, *7*, 131–153.

Kargupta, H., & Goldberg, D. E. (1994). *Decision making in genetic algorithms: A signal-to-noise perspective* (Tech. Rep. No. IllIGAL 94004). Urbana, IL: Illinois Genetic Algorithms Laboratory, University of Illinois at Urbana-Champaign.

Kauffman, S. A. (1993). *The origins of order.* Oxford: Oxford University Press.

Paredis, J. (1996). Symbiotic coevolution for epistatic problems. In W. Wahlster (Ed.), *12th European Conference on Artificial Intelligence.* Chichester: John Wiley & Sons, Ltd.

Potter, M. A. (1997). *The design and analysis of a computational model of cooperative coevolution.* Unpublished doctoral dissertation, George Mason University.

Potter, M. A., & De Jong, K. (1994). A cooperative coevolutionary approach to function optimization. In Y. Davidor, H.-P. Schwefel, & R. Männer (Eds.), *Parallel Problem Solving from Nature – PPSN III* (pp. 249–257). Berlin: Springer. (Lecture Notes in Computer Science 866)

Potter, M. A., & De Jong, K. (1995). Evolving neural networks with collaborative species. In *Proc. of the 1995 Summer Computer Simulation Conf.* (pp. 340–345). The Society of Computer Simulation.

Potter, M. A., & De Jong, K. (1998). The coevolution of antibodies for concept learning. In A. E. Eiben, T. Bäck, M. Schoenauer, & H.-P. Schwefel (Eds.), *Parallel Problem Solving from Nature – PPSN V* (pp. 530–539). Berlin: Springer. (Lecture Notes in Computer Science 1498)

Potter, M. A., De Jong, K., & Grefenstette, J. (1995). A coevolutionary approach to learning sequential decision rules. In L. J. Eshelman (Ed.), *Proc. of the Sixth Int. Conf. on Genetic Algorithms* (pp. 366–372). San Mateo, CA: Morgan Kaufman.

Salomon, R. (1996). Reevaluating genetic algorithm performance under coordinate rotation of benchmark functions. *BioSystems*, *39*(3), 263–278.

Schwefel, H.-P. (1977). *Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie* [Numeric optimization of computer models using the evolution strategy]. Basel: Birkhäuser. (Volume 26 of Interdisciplinary Systems Research)

Spedicato, E. (1975). *Computational experience with quasi-Newton algorithms for minimization problems of moderately large size* (Tech. Rep. No. CISE-N-175). Milano, Italy: Centro Informazioni Studi Espeienze.